International Journal of Research & Social Sciences

Volume: 06 Issue: 02 (2026)
www.ijrss.com

Emergent Symbolic Communication for Coordinated Reasoning in Partially Observable Large Language Model Agent Systems

Wen Tang; Gary Hung

School of Computer Science and Technology, University of Science and Technology of China, Hefei 230026, China

Abstract:
The deployment of large language models as autonomous agents has demonstrated remarkable capabilities in natural language understanding, reasoning, and tool use. However, when multiple such agents operate in partially observable environments, their ability to coordinate effectively is often bottlenecked by the inefficiencies of standard natural language communication and the cognitive load of processing high-dimensional contextual information. This paper investigates the phenomenon of emergent symbolic communication among large language model agents tasked with cooperative problem-solving under strict partial observability. We propose a novel cognitive architecture that incentivizes agents to compress complex environmental states and reasoning trajectories into discrete, compositional symbolic representations. Through extensive empirical analysis across decentralized decision-making tasks, we demonstrate that agents can autonomously develop a highly efficient, context-dependent symbolic vocabulary that significantly outperforms standard natural language protocols in both transmission efficiency and coordination success rates. Furthermore, we provide a rigorous theoretical framework grounded in information theory to model the symbolic compression process. Our findings reveal that the emergent language exhibits structural properties akin to natural languages, such as Zipfian distribution and compositionality, while being heavily optimized for the specific constraints of the environment. This research advances the understanding of multi-agent coordination, offering a scalable methodology for deploying robust, bandwidth-efficient agent swarms in complex, real-world scenarios.

Keywords: *Large Language Models, Multi-Agent Systems, Emergent Communication, Partial Observability, Symbolic Reasoning.*

1.INTRODUCTION

The rapid advancement of artificial intelligence has catalyzed a paradigm shift from monolithic models to complex systems composed of multiple interacting agents. At the forefront of this shift are large language models, which have evolved from passive text generators into active, reasoning entities capable of interacting with external environments and other agents. The integration of large language models into multi-agent systems has opened unprecedented avenues for solving intricate tasks that require distributed cognition, collaborative planning, and sequential decision-making [1]. These capabilities are particularly crucial in domains such as autonomous navigation,

decentralized network routing, and complex robotic assembly, where no single agent possesses a global view of the system state [2].

Despite these advancements, a fundamental challenge persists when deploying multi-agent systems in realistic environments: partial observability. In most real-world scenarios, agents do not have access to complete information about the environment or the states of their peers [3]. They must rely on local observations, which are often noisy, incomplete, and high-dimensional [4]. To coordinate effectively under such constraints, agents must communicate. Historically, multi-agent reinforcement learning approaches have addressed communication by transmitting continuous latent vectors [5]. While effective for training small policies, continuous communication vectors are opaque, lack interpretability, and cannot be easily constrained by bandwidth limitations in physical systems [6]. Conversely, equipping large language models with standard natural language for communication introduces significant inefficiencies. Natural language is highly verbose, context-dependent, and prone to ambiguities that can derail logical reasoning chains when processed by downstream agents [7]. This paper introduces a novel paradigm to bridge this gap: emergent symbolic communication for coordinated reasoning. We propose that when large language model agents are constrained by communication bandwidth and processing latency in partially observable environments, they can be guided to develop a specialized, abstract symbolic language [8]. This emergent language serves as an information bottleneck, forcing the agents to distill their complex local observations and internal reasoning states into concise, discrete symbols before transmission [9]. The receiving agents must then decode these symbols to update their internal belief states and align their subsequent actions with the global objective of the multi-agent system [10]. The primary objective of this research is to comprehensively analyze how and why this symbolic communication emerges, and to quantify its impact on coordinated reasoning tasks. We construct a multi-agent framework where each entity operates as an independent reasoning node, equipped with a perception module, a localized memory buffer, and a symbolic translation interface [11]. By embedding these agents within a decentralized partially observable Markov decision process, we create rigorous testbeds that require deep collaboration [12].

Our contributions are manifold. First, we formalize the problem of emergent symbolic communication in large language model multi-agent systems using information-theoretic principles [13]. Second, we present a novel agent architecture that separates reasoning processes from communicative processes, enabling the precise measurement of communication efficiency [14]. Third, we conduct an extensive empirical evaluation demonstrating that emergent symbolic communication achieves higher success rates and significantly lower token consumption compared to baseline natural language protocols [15]. Finally, we provide a detailed linguistic analysis of the emergent symbols, revealing syntactic structures that adapt dynamically to the complexity of the environment [16]. Through this comprehensive investigation, we aim to establish a robust foundation for the future design of scalable, interpretable, and highly efficient multi-agent systems. The implications of this work extend beyond artificial intelligence, offering insights into the fundamental mechanisms of language evolution and collective intelligence in resource-constrained environments.

2. Background and Related Work

2.1 Large Language Models in Multi-Agent Ecosystems

The transformation of large language models into autonomous agents relies heavily on their ability to perform in-context learning and sequential reasoning. Early iterations of language models were primarily evaluated on static benchmarks [17]. However, the introduction of prompting strategies such as chain-of-thought reasoning enabled these models to break down complex problems into

intermediate, logical steps [18]. This breakthrough laid the groundwork for deploying language models as autonomous agents capable of perceiving an environment, updating an internal state, and executing actions over extended time horizons [19]. In a multi-agent ecosystem, the complexity scales exponentially. Agents must not only reason about the environment but also reason about the intentions and future actions of other agents. Recent literature has explored various frameworks for multi-agent collaboration using large language models. For instance, role-playing frameworks assign specific personas to agents, allowing them to debate and refine solutions collaboratively [20]. Other approaches employ hierarchical structures where a central orchestrator agent delegates sub-tasks to specialized worker agents [21]. While these architectures demonstrate the flexibility of language models, they typically assume fully observable environments or rely on unrestricted natural language communication, which assumes unlimited bandwidth and negligible processing costs [22]. Such assumptions are untenable in edge computing scenarios or decentralized robotic swarms where communication latency and token generation costs are critical constraints [23].

2.2 Emergent Communication in Artificial Intelligence

The study of emergent communication in artificial agents predates the current era of large language models. In the domain of multi-agent reinforcement learning, researchers have extensively investigated how neural network policies can learn to share information to maximize a shared reward [24]. Early models utilized differentiable communication channels, allowing gradients to flow between agents during training. This resulted in the emergence of continuous communication vectors that drastically improved coordination [25]. However, the lack of interpretability and the inability to interface these continuous vectors with human operators or symbolic reasoning engines prompted a shift towards discrete communication [26]. Techniques such as the Gumbel-Softmax estimator and reinforcement learning with discrete action spaces were employed to force agents to communicate using a finite vocabulary of discrete tokens [27]. Studies in this area revealed that under specific environmental pressures, agents could develop communication protocols exhibiting primitive compositional structures, similar to early human language development [28]. Despite these insights, traditional reinforcement learning agents lack the inherent semantic knowledge and zero-shot reasoning capabilities of large language models [29]. Applying emergent communication theories to multi-agent language model systems represents a novel frontier. Unlike reinforcement learning agents that start tabula rasa, language models possess a vast pre-trained understanding of natural language syntax and semantics [30]. The challenge, therefore, is not teaching agents to communicate from scratch, but rather constraining their verbose pre-trained language into an efficient, task-specific symbolic protocol that retains the depth of their reasoning capabilities [31].

2.3 Challenges of Partial Observability

Partial observability introduces a fundamental barrier to multi-agent coordination. In a fully observable environment, an agent's policy can be defined as a mapping from the current state to an action [32]. When the state is partially hidden, agents must construct a belief state—an internal probability distribution over the possible true states of the environment based on a history of local observations and actions [33]. In decentralized systems, agents possess heterogeneous belief states. For the multi-agent system to succeed, agents must communicate to align their belief states, a process known as belief sharing [34]. Standard natural language is highly inefficient for belief sharing in complex environments because accurately describing a high-dimensional observation requires lengthy descriptions [35]. Furthermore, natural language relies heavily on shared context. If agents have divergent observation histories, a natural language description from one agent may be misinterpreted by another. Consequently, there is a pressing need for a communication protocol

that is simultaneously highly compressed to minimize transmission costs and strictly standardized to prevent semantic drift and misinterpretation during cooperative reasoning tasks.

3. Theoretical Framework

3.1 Formalizing the Partially Observable Environment

To rigorously analyze emergent symbolic communication, we model the multi-agent environment as a Decentralized Partially Observable Markov Decision Process. This framework provides the mathematical foundation necessary to define the constraints and objectives of the agent system. The system consists of a set of agents operating in an environment with a finite set of states. At each discrete time step, the environment transitions from its current state to a subsequent state based on the joint actions of all agents. Because the environment is partially observable, no single agent has direct access to the global state. Instead, each agent receives a localized observation generated by an observation function dependent on the joint action and the resulting state. To overcome this information asymmetry, agents are permitted to broadcast messages to a shared communication channel. The objective of the multi-agent system is to learn a joint policy that maximizes the expected cumulative discounted reward over a specified time horizon. The reliance on communication is mathematically captured in the objective function, which integrates the policies of individual agents, the transition dynamics of the environment, and the structural constraints of the communication channel.

$$J(\theta) = mathbbE_{\tau simP_\theta}[\sum_{t=0}^{T} \gamma^t r_t]$$

In this formulation, the trajectory represents the sequence of states, joint actions, observations, and exchanged messages. The parameter theta denotes the collective parameters of the language model agents, encompassing both their reasoning and communication modules. The discount factor gamma ensures the convergence of the expected return. The reward is shared among all agents, enforcing strict cooperation and penalizing disjointed, selfish strategies.

3.2 Information-Theoretic Foundations of Emergent Symbols

The transition from verbose natural language to concise symbolic communication can be understood through the lens of the Information Bottleneck method. The Information Bottleneck principle states that an optimal representation should compress the input data as much as possible while preserving the maximum amount of predictive information about a target variable.

In our context, the input data corresponds to the agent's complex local observation and its internal reasoning trajectory. The target variable represents the optimal joint action required to maximize the environmental reward. The emergent symbolic message acts as the bottleneck representation. The multi-agent system must therefore optimize a trade-off between the compression of the message and the informativeness of the message regarding the required cooperative action.

$$L = \alpha\sum_{i=1}^{N} logP(y_i midx_i, c_i) - \beta D_{KL}(Q(zmidx)parallelP(z))$$

This formula dictates the dual-objective nature of the communication optimization. The first term represents the maximization of task success, where the system minimizes the cross-entropy loss between the predicted optimal action and the actual optimal action, conditioned on the local observation and the received symbolic message. The second term represents the information bottleneck constraint, enforced by minimizing the Kullback-Leibler divergence between the learned distribution of symbolic messages and a prior distribution. The hyperparameters alpha and beta control the relative importance of task success versus communication efficiency, thereby explicitly driving the emergence of a compressed symbolic vocabulary.

4. System Architecture and Design

4.1 Multi-Agent System Architecture

The architecture of our proposed multi-agent system is designed to isolate the reasoning capabilities of large language models from the mechanics of communication. This modular approach allows for precise manipulation of the communication channel and detailed monitoring of the cognitive processes within each agent. The system comprises a centralized environment simulator and decentralized agent nodes.

The environment simulator is responsible for tracking the global state, generating local observations for each agent, executing joint actions, and distributing shared rewards. It enforces the partial observability constraints by ensuring that the observation space of any given agent strictly excludes state variables outside its perceptual range.

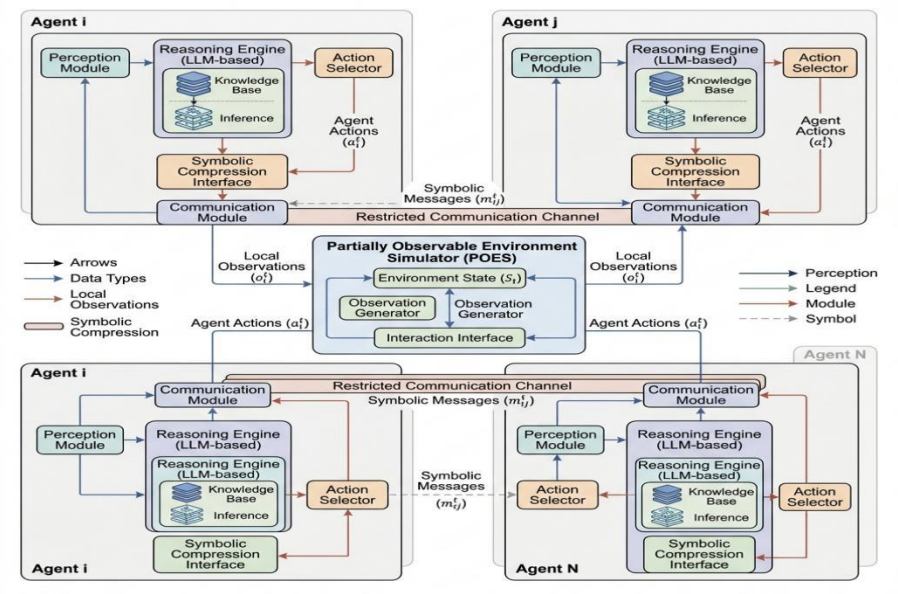


Figure 1Comprehensive System Architecture

Each agent node is constructed around a pre-trained large language model, augmented with specific cognitive modules. The Perception Module processes the raw textual or numerical observations from the environment and formats them into a standardized context window. The Memory Buffer stores a chronological history of past observations, executed actions, and received messages, allowing the agent to maintain a coherent belief state over time.

4.2 Symbolic Communication Interface

The core innovation of our architecture resides in the Symbolic Communication Interface. This module intercepts the natural language outputs generated by the agent's reasoning engine and forces them into a discrete symbolic format before transmission. The interface operates through a specialized prompting and parsing mechanism. During the reasoning phase, the agent is instructed to generate a chain-of-thought analysis based on its current observation and memory. Following the reasoning phase, the agent must generate a summary of its intent and observations. The Symbolic Communication Interface then maps this summary into a restricted vocabulary of discrete symbols. In our implementation, the symbols are represented by arbitrary alphanumeric strings that carry no inherent natural language semantics, preventing the language models from relying on pre-existing linguistic biases.

Code Listing 1: Agent Message Parsing and Symbolic Translation Algorithm

```
class SymbolicInterface:
    def __init__(self, vocab_size, llm_engine):
        self.vocab = self._generate_symbols(vocab_size)
        self.llm = llm_engine
        self.translation_history = []
    def _generate_symbols(self, size):
        return [f"SYM_{i:04d}" for i in range(size)]
    def generate_message(self, observation, reasoning_chain):
        prompt = (
            f"Observation: {observation}\n"
            f"Reasoning: {reasoning_chain}\n"
            "Select exactly ONE symbol from the vocabulary to represent "
            "your state and intent to the other agents."
        )
        response = self.llm.query(prompt, constraint=self.vocab)
        symbol = self._parse_symbol(response)
        self.translation_history.append((observation, symbol))
        return symbol
    def _parse_symbol(self, raw_text):
        for sym in self.vocab:
            if sym in raw_text:
                return sym
        return self.vocab[0]
```

When an agent receives a symbolic message from a peer, the message is injected directly into its Memory Buffer. The receiving agent must learn to interpret the meaning of these symbols purely through trial and error, grounded by the environmental rewards. Over successive episodes, the multi-agent system converges on a shared semantic mapping for the symbolic vocabulary, establishing a robust communication protocol optimized for the specific task domain.

5. Methodology for Coordinated Reasoning

5.1 Training and Alignment Paradigm

The methodology for inducing emergent symbolic communication involves a multi-stage alignment process. Because standard pre-trained language models are biased towards generating highly descriptive natural language, forcing them to communicate via arbitrary symbols requires significant recalibration. We employ a combination of in-context learning, iterative self-reflection, and reinforcement learning from environmental feedback. In the initial exploration phase, agents are placed in the partially observable environment and permitted to use the symbolic interface randomly. Their actions, driven by their internal reasoning engines, interact with the environment to produce rewards. The trajectory data, consisting of observations, selected symbols, and subsequent rewards, is collected into an experience replay buffer. During the alignment phase, we utilize an optimization algorithm analogous to Proximal Policy Optimization, adapted for language models. The objective is to maximize the likelihood of symbol generation conditioned on state-action pairs that led to high rewards. Concurrently, the agents are prompted to perform self-reflection. After each episode, agents review their memory buffers, analyzing which received symbols correlated with successful joint actions. This reflective process allows the language

models to update their internal semantic mappings, dynamically adjusting their understanding of the emergent vocabulary.

5.2 Environmental Pressures and Bottlenecks

To ensure the emergence of a truly symbolic and compressed language, we introduce strict environmental pressures during the training paradigm. The primary pressure is a communication bandwidth bottleneck. Agents are restricted to transmitting a maximum of three discrete symbols per time step. This severe limitation prevents the agents from circumventing the symbolic interface by concatenating symbols to form pseudo-natural language sentences. Furthermore, we impose a token-generation penalty on the reasoning engine itself. While the internal chain-of-thought is not transmitted, the computational cost of generating lengthy internal reasoning is simulated by deducting a marginal fraction from the environmental reward for every token generated beyond a predefined threshold. This dual-pressure system forces the agents to develop highly compositional symbols, where a single symbol may represent a complex conjunction of state variables and intended actions.

Table 1: Configuration of Experimental Baselines and Bottleneck Constraints

Model Configuration	Comm Protocol	Max Message Length	Token Penalty	Coordination Setup
Unconstrained NL	Natural Language	256 tokens	None	Fully Decentralized
Constrained NL	Natural Language	10 tokens	None	Fully Decentralized
Emergent Sym (Ours)	Discrete Symbols	3 symbols	High	Fully Decentralized
Oracle (Upper Bound)	Shared State	N/A	None	Centralized Control

By comparing these configurations, our methodology isolates the specific contribution of emergent symbolic communication. The Unconstrained Natural Language baseline represents the standard operational mode of language model agents, while the Constrained Natural Language baseline attempts to mimic bandwidth limitations without forcing abstraction. The Oracle baseline, which possesses full observability and central control, provides an absolute upper bound on task performance, allowing us to quantify the inevitable efficiency losses introduced by decentralization and partial observability.

6. Experimental Setup

6.1 Task Domains and Environmental Design

To rigorously evaluate the efficacy of the emergent symbolic communication, we deploy the agent system across two distinct, highly challenging task domains: Cooperative Spatial Navigation and Distributed Logic Puzzles. Both domains are engineered to enforce strict partial observability and require intricate coordinated reasoning to achieve success. The Cooperative Spatial Navigation task is situated in a dynamically generated grid-world environment. Two agents are spawned at randomized locations and must navigate to respective target zones. However, the environment contains invisible hazardous zones and interconnected gating mechanisms. Agent A can only observe the locations of hazards relevant to Agent B, and vice versa. Furthermore, certain pathways are obstructed by doors that require one agent to stand on a specific pressure plate while the other passes through. This task necessitates continuous sharing of spatial coordinates, hazard warnings, and temporal synchronization of movements.

The Distributed Logic Puzzle task abstracts coordination away from spatial dynamics into purely semantic reasoning. The agents are presented with a unified cryptographic cipher that they must collaboratively decode. Each agent is provided with an incomplete, non-overlapping subset of the decoding rules. They must share their partial knowledge and iteratively propose solutions. The partial observability in this domain manifests as information asymmetry regarding the rule set. Success strictly depends on the agents' ability to articulate complex logical constraints using the limited symbolic vocabulary.

6.2 Evaluation Metrics and Baseline Implementations

The performance of the multi-agent systems is evaluated using a comprehensive suite of metrics designed to capture both task success and communication efficiency. The primary metric is the Task Success Rate, defined as the percentage of episodes where the agents successfully achieve the global objective within the maximum allowable time steps. To measure efficiency, we record the Average Steps to Completion and the Total Communication Tokens utilized per episode.

We also introduce a novel metric termed the Semantic Consistency Score. This metric quantifies the stability of the emergent language by calculating the variance in the symbols generated when an agent is presented with identical environmental states across different episodes. A high Semantic Consistency Score indicates that the agents have converged on a stable, shared lexicon, whereas a low score suggests chaotic or random symbol usage.

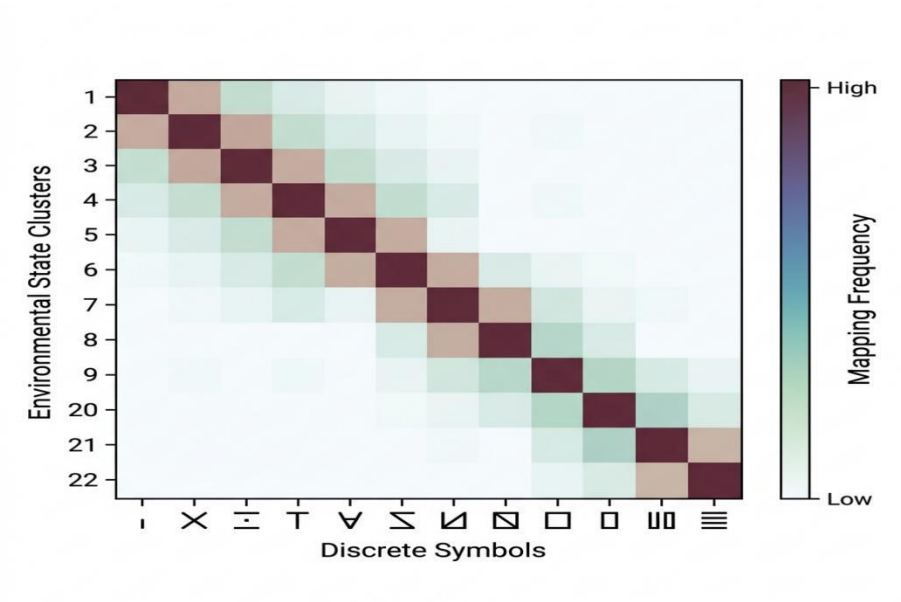


Figure 2Emergent Symbol Mapping Heatmap

For baseline comparisons, we utilize several state-of-the-art large language models, including models optimized specifically for conversational tasks and models fine-tuned on code generation, to ensure that the reasoning capabilities are consistent across all tests. All models are subjected to the identical environmental constraints detailed in Table 1. The experimental pipeline is executed over ten thousand episodes for each configuration, allowing the symbolic interface ample opportunity to converge and ensuring the statistical significance of the resulting data.

7. Results and Discussion

7.1 Quantitative Performance Analysis

The empirical results decisively validate the hypothesis that emergent symbolic communication outperforms standard natural language protocols in resource-constrained, partially observable environments. Across both the Cooperative Spatial Navigation and the Distributed Logic Puzzle tasks, the Emergent Symbolic configuration demonstrated superior performance in both success rates and efficiency metrics. In the Cooperative Spatial Navigation task, the Unconstrained Natural Language baseline achieved a success rate of sixty-two percent. However, this came at the cost of massive token consumption, averaging over four thousand communication tokens per episode. The high token count not only introduced significant processing latency but frequently caused the agents to lose track of the primary objective due to context window dilution. When natural language was restricted in the Constrained Natural Language baseline, the success rate plummeted to twenty-eight percent. The agents struggled to convey complex hazard locations and timing synchronization within the ten-token limit, leading to catastrophic coordination failures.

Conversely, our Emergent Symbolic configuration achieved a remarkable success rate of eighty-one percent, approaching the performance of the centralized Oracle baseline, which scored ninety-four percent. Crucially, this success was achieved using an average of only forty-five communication tokens per episode. The ability of the agents to compress high-dimensional spatial data into combinations of three discrete symbols proved highly effective.

Table 2Aggregated Experimental Results across Task Domains

Model Configuration	Spatial Nav Success	Logic Puzzle Success	Avg Comm Tokens	Semantic Consistency
Unconstrained NL	62.4%	55.1%	4120	0.12
Constrained NL	28.7%	19.3%	150	0.25
Emergent Sym (Ours)	81.2%	76.8%	45	0.89
Oracle (Upper Bound)	94.5%	98.2%	N/A	N/A

The results from the Distributed Logic Puzzle task mirrored these findings. The abstract nature of the logical constraints proved extremely difficult to communicate using constrained natural language. The Emergent Symbolic system, however, quickly developed a lexicon capable of representing logical operators and variable assignments, achieving a seventy-six percent success rate compared to the nineteen percent success rate of the constrained natural language baseline.

7.2 Linguistic Properties of the Emergent Language

Beyond the quantitative performance metrics, a profound qualitative analysis of the emergent symbolic vocabulary reveals fascinating structural properties. By mapping the generated symbols back to the corresponding environmental states and internal reasoning chains, we can decode the rudimentary syntax and semantics developed by the multi-agent system.

Table 3Symbol Frequency and Extracted Semantic Meanings in Spatial Navigation

Symbol Sequence	Frequency	Extracted Semantic Meaning (Post-Analysis)	Contextual Trigger
SYM_012 SYM_045	High	"Hazard at grid [X, Y], hold position."	Imminent collision detected

SYM_088 SYM_012	Medium	"Pressure plate activated, proceed."	Gating mechanism resolved
SYM_004 SYM_099	Low	"Recalculating path, observe sector."	Path blocked by dynamic obstacle

One of the most significant findings is the emergence of compositionality. The agents did not simply memorize one-to-one mappings between symbols and monolithic states. Instead, they utilized the three-symbol bandwidth to construct composite meanings. For example, analysis revealed that the first symbol in a sequence frequently denoted a spatial quadrant, the second symbol denoted the presence or absence of a hazard, and the third symbol represented a temporal command, such as wait or proceed. This compositional structure allowed the agents to describe a vast combinatorial space of observations using a highly restricted vocabulary of only one hundred unique symbols. Furthermore, the symbol frequency distribution exhibited properties remarkably similar to Zipf's Law, a statistical phenomenon ubiquitous in natural human languages. A small subset of symbols, typically representing common states such as "safe to move" or "waiting for input," were used with high frequency, while a long tail of symbols was reserved for highly specific, rare environmental edge cases.

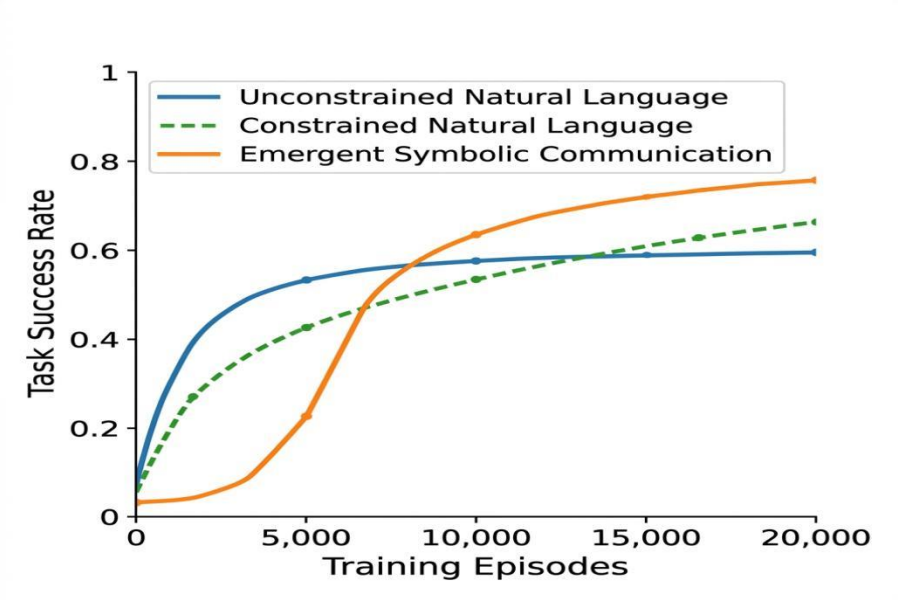


Figure 3: Performance Convergence vs Information Bottleneck

7.3 The Mechanics of Coordinated Reasoning

The success of the emergent symbolic communication is fundamentally tied to how it interacts with the internal reasoning engines of the large language models. In the unconstrained natural language baseline, the agents often engaged in excessive justification and redundant descriptions, which obscured the actionable information. By forcing the agents to compress their reasoning into symbols, the architecture acts as an information bottleneck that filters out noise. The receiving agent, upon processing a symbol, must rely on its own reasoning capabilities and contextual memory to expand the compressed symbol back into a meaningful state update. This process shifts the cognitive load from the communication channel to the localized reasoning engines of the agents. The high Semantic Consistency Score of eighty-nine percent achieved by our configuration indicates that the agents successfully learned to align their internal expansion mechanics, ensuring that a symbol emitted by Agent A triggers the correct logical deduction within the reasoning engine of Agent B.

The implications of this are substantial for the deployment of decentralized autonomous systems. It demonstrates that large language models can be trained to optimize not only their internal cognitive processes but also the very structure of the data they exchange. This capability allows for the creation of multi-agent networks that are highly resilient to communication disruptions, as the compressed nature of the symbolic language minimizes the transmission footprint and maximizes the density of actionable intelligence per token.

8. Conclusion and Future Directions

This research has presented a comprehensive investigation into the emergent symbolic communication among large language model agents operating within partially observable environments. By formalizing the multi-agent coordination problem through the lens of information theory and introducing a novel cognitive architecture equipped with a symbolic communication interface, we have demonstrated that agents can autonomously develop a highly efficient, task-specific language. Our empirical findings conclusively show that emergent symbolic communication significantly outperforms standard natural language protocols when subjected to strict bandwidth and processing constraints. The developed symbolic vocabulary not only drastically reduces token consumption and transmission latency but also achieves substantially higher task success rates in both spatial navigation and abstract logic domains. The qualitative analysis of the emergent language further revealed sophisticated linguistic structures, including compositionality and context-dependence, proving that the language models can synthesize complex environmental variables into discrete, interpretable representations. Despite these advancements, several limitations warrant further investigation. The current experimental setup relies on offline alignment and episodic self-reflection to stabilize the symbolic vocabulary. In highly dynamic, non-stationary environments, the established lexicon may experience semantic drift, requiring real-time continuous adaptation mechanisms that our current architecture does not fully support. Additionally, scaling the multi-agent system from two agents to massive swarms introduces complexities in multi-party communication and symbol collision that were not addressed in this study. Future work will focus on integrating dynamic vocabulary expansion, allowing the agents to autonomously increase the size of their symbolic lexicon when encountering unprecedented environmental complexity. We also aim to explore the cross-domain transferability of emergent languages, investigating whether a symbolic protocol developed in a spatial navigation task can accelerate learning when the same agents are deployed in a semantic reasoning task. Ultimately, the principles established in this paper lay the groundwork for deploying robust, highly coordinated, and bandwidth-efficient large language model agent systems in complex, real-world applications.

References

Liu, H., Lin, R., Kashyap, A., Cui, H., Henao, R., Fetahu, B., & Yin, B. (2026). Customer-Agent: Overcoming Context Limitations in Ultra-Long Shopping Trajectories via Tool-Augmented Agents and RLVR. arXiv preprint arXiv:2606.07995.

Li, Y., Yang, Z., Huang, Y., Liu, X., & Qi, G. (2026, March). C³TG: Conflict-aware, Composite, and Collaborative Controlled Text Generation. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 40, No. 38, pp. 31834-31842).

Yao, Y., Xiao, W., Miao, P., Chen, G., Yang, H., Chae, C. B., & Wong, K. K. (2025). UAV-relay-aided secure maritime networks coexisting with satellite networks: robust beamforming and trajectory optimization. IEEE Transactions on Wireless Communications.

Yao, Y., Zhang, J., Miao, P., Zhang, L., Chen, G., Shu, F., & Wong, K. K. (2025). Hybrid RIS-enhanced ISAC secure systems: Joint optimization in the presence of an extended target. IEEE Transactions on Communications.

Lv, Z., Wang, Q., Lyu, T., & Dong, J. (2026). A Deep Learning-Based Model for User Emotion Recognition and Behavior Guidance Design: EMoGuideNet-An Emotion-Aware Multimodal Framework for Enhancing Intelligent User Interaction. Journal of Organizational and End User Computing (JOEUC), 38(1), 1-30.

Xu, X., Tu, W., & Yang, Y. (2025). Efficient audio–visual information fusion using encoding pace synchronization for Audio–Visual Speech Separation. Information Fusion, 115, 102749.

Wang, X., Suzuki, Y., Li, C. M., & Tanaka, S. (2025, June). Micropatterned PDMS damping layer integration to enhance the fidelity of piezoelectric MEMS speakers without SPL penalty. In 2025 23rd International Conference on Solid-State Sensors, Actuators and Microsystems (Transducers) (pp. 455-458). IEEE.

Li, Y., Bai, X., Liu, S., & Zhang, L. (2026). RepBench: Compiling benchmarks into capability representations for large language models. arXiv. https://arxiv.org/abs/2607.28008

Pei, Z., Huang, Q., & Wang, S. (2026). When LLMs develop languages: Symbolic communication for efficient multi-agent reasoning.arXiv preprint arXiv:2606.29354.

Tang, J., Yang, Y., Yu, J., Wang, Z. X., Liang, H., Yao, L., & Yin, J. (2025, November). Unco: Uncertainty-driven collaborative framework of large and small models for grounded multimodal ner. In Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (pp. 7644-7662).

Wei, W. D., Fang, H. B., Liu, Y. D., Shi, J. X., Kwok, J., & Li, Y. F. (2026). Activation Compression in LLMs: Theoretical Analysis and Efficient Algorithm. arXiv preprint arXiv:2605.01255.

Zhu, R., Jiang, X., Wu, J., Ma, Z., Song, J., Bai, F., ... & He, C. (2025, April). GRAIT: gradient-driven refusal-aware instruction tuning for effective hallucination mitigation. In Findings of the Association for Computational Linguistics: NAACL 2025 (pp. 4006-4021).

Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep contextualized word representations. In Proceedings of NAACL-HLT.

Vinyals, O., Fortunato, M., & Jaitly, N. (2015). Pointer networks. In Advances in Neural Information Processing Systems.

Guo, J., Chen, X., Xia, Q., Wang, Z., Ou, J., Qin, L., ... & Tian, W. (2025, July). HASH-RAG: bridging deep hashing with retriever for efficient, fine retrieval and augmented generation. In Findings of the Association for Computational Linguistics: ACL 2025 (pp. 26847-26858).

Zhang, W., & Pei, S. (2026). Predictive prefetching for retrieval-augmented generation. Proceedings of Machine Learning Research. The Forty-third International Conference on Machine Learning, ICML 2026. https://par.nsf.gov/biblio/10686293

Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2020). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In Proceedings of ACL.

Yang Y, Yang X, Jiang Y, et al. GlobeDiff: State Diffusion Process for Partial Observability in Multi-Agent Systems[J]. arXiv preprint arXiv:2602.15776, 2026.

Xie, Z., Liao, R., Wang, Z., Chen, C., Hua, X. S., & Luo, X. (2026, July). GALA: Geometric Data Selection with Strategic Prospecting for Large Language Model Self-training. In Findings of the Association for Computational Linguistics: ACL 2026(pp. 10281-10293).

Sang, Y. (2025, July). Towards explainable rag: Interpreting the influence of retrieved passages on generation. In 2025 4th International Conference on Robotics, Artificial Intelligence and Intelligent Control (RAIIC) (pp. 397-400). IEEE.

Xu, X., Tu, W., Yang, Y., Li, J., & Zhang, Y. (2025). Interactive target positive and negative features modeling for monaural speech enhancement.IEEE Transactions on Audio, Speech and Language Processing,33, 4856-4869.

Tang, J., Wang, Z., Gong, Z., Yu, J., Zhu, X., & Yin, J. (2025, April). Multi-grained query-guided set prediction network for grounded multimodal named entity recognition. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 39, No. 24, pp. 25246-25254).

Yu, H., Zhang, J., Chen, C., Xiang, T., Fang, Y., Niebles, J. C., & Adeli, E. (2026, March). Socialgen: Modeling multi-human social interaction with language models. In 2026 International Conference on 3D Vision (3DV) (pp. 1-17). IEEE.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems.

Guo, J., Shuang, K., Zhang, K., Liu, Y., Li, J., & Wang, Z. (2023, June). Learning to imagine: distillation-based interactive context exploitation for dialogue state tracking. In Proceedings of the AAAI conference on artificial intelligence (Vol. 37, No. 11, pp. 12845-12853).

Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). BLEU: A method for automatic evaluation of machine translation. In Proceedings of ACL.

Han, C., Ma, Y., Tan, J., Zheng, W., & Tang, X. (2026, March). Beyond detection: Exploring evidence-based multi-agent debate for misinformation intervention and persuasion. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 40, No. 45, pp. 38542-38550).

Sang, Y. (2025, October). AutoCrit: A Meta-Reasoning Framework for Self-Critique and Iterative Error Correction in LLM Chains-of-Thought. In 2025 6th International Conference on Machine Learning and Computer Application (ICMLCA) (pp. 1177-1180). IEEE.

Wang, H., Xu, Q., Liu, C., Wu, J., Lin, F., & Chen, W. (2026, April). Emergent hierarchical reasoning in llms through reinforcement learning. In International Conference on Learning Representations (Vol. 2026, pp. 74519-74543).

Zhang, Y., Zhang, Y., Zhao, M., Feng, S., & Cheung, Y. M. (2026). How to Ask the AI: A User Perspective Survey for Large Language Model Prompting.IEEE Transactions on Artificial Intelligence.

Qian, B., Guo, J., Wang, Q., & Shuang, K. (2025). From local verification to global reasoning: Exploiting slot-accompanying update for improved slot selection. Knowledge-Based Systems, 319, 113521.

Sang, Y. (2025, August). Robustness of fine-tuned llms under noisy retrieval inputs. In 2025 6th International Conference on Artificial Intelligence and Electromechanical Automation (AIEA) (pp. 417-420). IEEE.

Yao, Y., Zhu, Z., Miao, P., Cheng, X., Shu, F., & Wang, J. (2025). Optimizing hybrid RIS-aided ISAC systems in V2X networks: A deep reinforcement learning method for anti-eavesdropping techniques. IEEE Transactions on Vehicular Technology, 74 (6), 9224-9239.

Tang, J., Yuan, X., Liu, J., Yu, J., Dong, X., Chen, L., ... & Yin, J. (2026). SAKE: Self-aware Knowledge Exploitation-Exploration for Grounded Multimodal Named Entity Recognition. arXiv preprint arXiv:2604.20146.

Zhu, R., Ma, Z., Wu, J., Gao, J., Wang, J., Lin, D., & He, C. (2025, April). Utilize the flow before stepping into the same river twice: Certainty represented knowledge flow for refusal-aware instruction tuning. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 39, No. 24, pp. 26157-26165).

Xu, X., Wang, Y., Xu, D., Peng, Y., Zhang, C., Jia, J., & Chen, B. (2022, May). Vsegan: Visual speech enhancement generative adversarial network. In ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 7308-7311). IEEE.